What is claimed is:

1. A substantially purified nucleic acid molecule expressed in response to polycyclic aromatic hydrocarbon exposure comprising:

(a) a nucleic acid molecule encoding a protein selected from SEQ ID NOs:7, 8 and 14;

(b) a nucleic acid molecule selected from the group consisting of SEQ ID NOs:1-5 and 9-13; or

(c) a nucleic acid molecule which is the complement of the nucleic acid molecule of (a) or (b) wherein each and every nucleotide of the complement is complementary to each and every nucleotide of the nucleic acid molecule of (a) or (b).

2. A substantially purified nucleic acid molecule expressed in response to polycyclic aromatic hydrocarbon exposure comprising:

(a) a nucleic acid molecule encoding a protein of SEQ ID NO:6; or

(b) a nucleic acid molecule which is the complement of the nucleic acid molecule of (a) wherein each and every nucleotide of the complement is complementary to each and every nucleotide of the nucleic acid molecule of (a).

3. A method of using a nucleic acid molecule to screen a library of molecules or compounds to identify at least one ligand which specifically binds the nucleic acid molecule, the method comprising:

(a) combining the nucleic acid molecule of claim 1 with a library of molecules or compounds under conditions to allow specific binding; and

(b) detecting specific binding, thereby identifying a ligand which specifically binds the nucleic acid molecule.

4. A method of using a nucleic acid molecule to screen a library of molecules or compounds to identify at least one ligand which specifically binds the nucleic acid molecule, the method comprising:

(a) combining the nucleic acid molecule of claim 2 with a library of molecules or compounds under conditions to allow specific binding; and

(b) detecting specific binding, thereby identifying a ligand which specifically binds the nucleic acid molecule.

5. The method of claim 3 wherein the library is selected from DNA molecules, RNA molecules, peptide nucleic acids, mimetics, and proteins.

6. The method of claim 4 wherein the library is selected from DNA molecules, RNA molecules, peptide nucleic acids, mimetics, and proteins.

7. A ligand identified by the method of claim 3 which modulates the activity of the nucleic acid

molecule.

8. A ligand identified by the method of claim 5 which modulates the activity of the nucleic acid molecule.

9. A method of using a nucleic acid molecule to purify a ligand which specifically binds the nucleic acid molecule, the method comprising:

(a) combining the nucleic acid molecule of claim 1 with a sample under conditions to allow specific binding;

(b) detecting specific binding between the nucleic acid molecule and a ligand;

(c) recovering the bound nucleic acid molecule; and

(d) separating the nucleic acid molecule from the ligand, thereby obtaining purified ligand.

10. A method of using a nucleic acid molecule to purify a ligand which specifically binds the nucleic acid molecule, the method comprising:

(a) combining the nucleic acid molecule of claim 2 with a sample under conditions to allow specific binding;

(b) detecting specific binding between the nucleic acid molecule and a ligand;

(c) recovering the bound nucleic acid molecule; and

(d) separating the nucleic acid molecule from the ligand, thereby obtaining purified ligand.

11. A method for diagnosing a disorder or condition associated with the altered expression of a gene expressed in response to polycyclic aromatic hydrocarbon exposure in a plurality of biological samples, the method comprising the steps of:

(a) hybridizing a nucleic acid molecule of claim 1 to a sample under conditions effective to form one or more hybridization complexes;

(b) detecting the hybridization complexes; and

(c) comparing the levels of the hybridization complexes with the level of hybridization complexes in a control sample, wherein the altered level of hybridization complexes compared with the level of hybridization complexes of a control sample indicates the presence of the disorder or condition.

12. A method for diagnosing a disorder or condition associated with the altered expression of a gene expressed in response to polycyclic aromatic hydrocarbon exposure in a plurality of biological samples, the method comprising the steps of:

(a) hybridizing a nucleic acid molecule of claim 2 to a sample under conditions effective to form one or more hybridization complexes;

(b) detecting the hybridization complexes; and

(c) comparing the levels of the hybridization complexes with the level of hybridization

complexes in a control sample, wherein the altered level of hybridization complexes compared with the level of hybridization complexes of a control sample indicates the presence of the disorder or condition.

13. A method for detecting or diagnosing effect of a compound on expression level of at least one nucleic acid molecule in a subject, the method comprising:

(a) treating the subject with the compound;

(b) obtaining a sample containing nucleic acid molecules from the subject;

(c) contacting the sample with at least one nucleic acid molecule of claim 1 under conditions for the formation of hybridization complexes; and

(d) detecting at least one hybridization complex, wherein the presence, absence, or change in amount of hybridization complex when compared with hybridization complex formed with a sample from an untreated subject indicates the effect of the compound.

14. A method for detecting or diagnosing effect of a compound on expression level of at least one nucleic acid molecule in a subject, the method comprising:

(a) treating the subject with the compound;

(b) obtaining a sample containing nucleic acid molecules from the subject;

(c) contacting the sample with at least one nucleic acid molecule of claim 2 under conditions for the formation of hybridization complexes; and

(d) detecting at least one hybridization complex, wherein the presence, absence, or change in amount of hybridization complex when compared with hybridization complex formed with a sample from an untreated subject indicates the effect of the compound.

15. A substantially purified protein expressed in response to polycyclic aromatic hydrocarbon exposure, comprising

(a) a protein selected from SEQ ID NOs:6-8 or a portion thereof; and

(b) an oligopeptide comprising at least 6 sequential amino acids of the protein of (a); and

(c) an immunogenic fragment of the protein of (a).

16. A protein of claim 15, comprising the amino acid sequence of SEQ ID NO:6.

17. A protein of claim 15, comprising the amino acid sequence of SEQ ID NO:7.

18. A protein of claim 15, comprising the amino acid sequence of SEQ ID NO:8.

19. A composition comprising a protein of claim 15 and a pharmaceutical carrier.

20. A method for using a protein to screen a library of molecules or compounds to identify at least one ligand which specifically binds the protein, the method comprising:

(a) combining the protein of claim 15 with the library of molecules or compounds under conditions to allow specific binding; and

(b) detecting specific binding between the protein and ligand, thereby identifying a ligand which specifically binds the protein.

21. The method of claim 20 wherein the library is selected from DNA molecules, RNA molecules, peptide nucleic acids, mimetics, proteins, agonists, antagonists, and antibodies.

22. A ligand identified by the method of claim 20 which modulates the activity of the protein.

23. A method of using the protein to purify a ligand from a sample, the method comprising:

(a) combining the protein of claim 15 with a sample under conditions to allow specific binding;

(b) detecting specific binding between the protein and a ligand;

(c) recovering the bound protein; and

(d) separating the protein from the ligand, thereby obtaining purified ligand.

24. An antibody which specifically binds to the protein of claim 15.

25. A diagnostic test for a condition or disease associated with the expression of a protein in a biological sample comprising the steps of:

(a) combining the biological sample with an antibody of claim 24, under conditions suitable for the antibody to bind the protein and form an antibody:protein complex; and

(b) detecting the complex, wherein the presence of the complex correlates with the presence of the protein in the biological sample.

26. The antibody of claim 24, wherein the antibody is:

(a) a chimeric antibody;

(b) a single chain antibody;

(c) a Fab fragment;

(d) a $F(ab')_2$ fragment, or

(e) a humanized antibody.

27. A composition comprising an antibody of claim 24 and an acceptable excipient.

28. A method of diagnosing a condition or disease associated with the expression of a protein in a subject, comprising administering to said subject an effective amount of the composition of claim 26.

29. A composition of claim 26, wherein the antibody is labeled.

30. A method of diagnosing a condition or disease associated with the expression of a protein in a subject, comprising administering to said subject an effective amount of the composition of claim 29.

31. A method of preparing a polyclonal antibody comprising:

(a) immunizing an animal with a protein of claim 15 under conditions to elicit an antibody

response;

(b) isolating antibodies from said animal; and

(c) screening the isolated antibodies with the protein thereby identifying a polyclonal antibody which binds specifically to a protein of SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8.

32. An antibody produced by a method of claim 31.

33. A composition comprising the antibody of claim 32 and a suitable carrier.

34. A method of making a monoclonal antibody comprising:

(a) immunizing an animal with a protein of claim 15 under conditions to elicit an antibody response;

(b) isolating antibody producing cells from the animal;

(c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

(d) culturing the hybridoma cells; and

(e) isolating from the culture monoclonal antibody which binds specifically to a protein of SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8.

35. A monoclonal antibody produced by a method of claim 34.

36. The antibody of claim 24, wherein the antibody is produced by screening a Fab expression library.

37. The antibody of claim 24, wherein the antibody is produced by screening a recombinant immunoglobulin library.

38. A method for detecting a protein in a sample comprising the steps of:

(a) incubating the antibody of claim 24 with a sample under conditions to allow specific binding of the antibody and the protein; and

(b) detecting specific binding, wherein specific binding indicates the presence of a protein of SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8 in the sample.

39. A method of purifying a protein from a sample, the method comprising:

(a) incubating the antibody of claim 24 with a sample under conditions to allow specific binding of the antibody and the protein; and

(b) separating the antibody from the sample and obtaining purified protein of SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8.

TABLE 1

Rat SEQ ID NO	Rat Incyte ID No.	Human SEQ ID NO.	Human Incyte ID No.	Unique Fragment	Identity	Human Tissue Expression Profile
9	700062690	1	1851405	1587-1634	78.3%	27% reproductive tissue, 15% nervous tissue, 14% cardiovascular, and 14% gastrointestinal tissues
10	700140450	2	1991226	109-131	62.1%	50% hematopoietic/immune tissue
11	700367683	3	253053	690-740	61.3%	31% reproductive tissue and 25% nervous tissue
12	701316810	4	2009569	327-368	95.4%	67% reproductive tissue
13	700139271	5	1642580	196-231	49.3%	44% liver tissue and 22% hematopoietic/immune tissue

<110> Kaser, Matthew R.
      Azimzai, Yalda
      Yue, Henry

<120> POLYCYCLIC AROMATIC HYDROCARBON INDUCED MOLECULES

<130> PB-0011 US

<140> To Be Assigned
<141> Herewith

<160> 14

<170> PERL Program

<210> 1
<211> 2439
<212> DNA
<213> Homo sapiens

<220>
<221>
<223> 1851405

<300>

<400> 1
```
cttcagttcc tcgaagagat tcactttcaa aaacatcaac tccttataac aaatcaaaca   60
aagcagcaag ccaacaaggg accccatggg aaacacttgt cgtgtttgct atcaacttga  120
agcaattaaa cgttcaaatg aatatgagta atgtaatggg aaatacaact tggacaacta  180
gtggtttgaa gagccagggc cgtctgtcag taggaagtaa tcgtgatcga gagatcagca  240
tgtctgttgg tctgggaaga tcacaattag attctaaagg aggagtagtt ggagggacca  300
tagatgtcaa tgctttggag atggttgctc atatttctga acatccaaat cagcaaccca  360
gtcacaaaat tcagattact atgggttcta ctgaagctcg tgttgattac atgggctcaa  420
gtatcctcat gggcatcttc agtaatgctg atcttaagct tcaggatgaa tggaaagtaa  480
acttgtataa tacattggat tcaagcataa ctgataaaag tgagattttc gtccatggag  540
atttgaagtg ggatattttc caagtaatga tatcaaggtc aaccacacca gatctgataa  600
aaataggaat gaagctccag gaatttttca cacaacaatt tgataccagc aaacgagctc  660
tgtctacctg gggaccagtt ccttaccttc cgccaaagac aatgactagc aacctagaaa  720
aaagttcaca agaacaatta cttgatgcag cacatcatcg acactggcct ggagtattga  780
aggtggtatc aggatgccac atatccttat ttcagattcc attaccagaa gatggaatgc  840
aatttggagg atcaatgagc ttacatggaa atcatatgac actggcatgt tttcatggtc  900
caaattttcg ttcaaaatct tgggcccttt ttcatttaga agaaccaaat attgcttttt  960
ggactgaagc tcagaaaatc tgggaagatg gctccagtga tcattctaca tatattgtac 1020
aaacactaga ttttcacctg ggtcataata ctatggttac caaaccatgt ggtgctttgg 1080
aaagtcctat ggcaacaata accaagataa caaggcgtcg ccatgaaaat ccaccccatg 1140
gagtagcaag tgtgaaagaa tggttcaatt atgttacagc tacaaggaat gaagagctaa 1200
atctgcttcg taatgttgat gctaacaaca ctgagaatag cactactgtg aagaattcta 1260
gtttgttgag tggattcaga ggaggttcta gctacaacca tgaaacagag actatctttg 1320
cattaccaag gatgcagctt gactttaaat ccattcatgt tcaagaacca caggagcctt 1380
cattacagga tgccagcctg aagccaaaag tagaatgtag tgtggtgaca gagttcactg 1440
accacatttg tgtgactatg gatgctgagc tcatcatgtt tcttcatgat ttagtatcag 1500
cttatcttaa agaaaaagaa aaagccatct ttccacctcg gattttatct actcgaccag 1560
gacaaaaaag tccaattatt atacatgacg acaattcctc tgataaagat agagaagata 1620
gcatcactta tactactgtg gactggagag attttatgtg caatacatgg cacctagaac 1680
ctactcttag attaatttct tggactggaa gaaagattga tccagtaggt gttgattata 1740
ttcttcaaaa attgggcttt catcatgcta ggactactat tcctaaatgg cttcaaagag 1800
gagtcatgga tccactggac aaggttctgt cagttcttat caaaaagctc ggtactgcac 1860
tacaggatga aaaggaaaag aaaggcaaag acaaagaaga acactaaaaa agtaatttga 1920
tctgtgaaca aattatgatt gtgtctgttt tattacactg gagtgttttt ttagtataat 1980
aatttgaaat ataactttaa aataattcta aatttgtggc tataattaaa agtttgtaag 2040
ttaacctgtt ctagttccat cattctgtgt acagtgaagt attgcatgat aatgtaaatt 2100
ttgtgaaaaa ctagattaaa atatataact gcttgttatg gtttataatt atataatgtg 2160
caatacaatt cctgcatctt taaaatgtct gcagaataac tgtgaatttt tttgttattg 2220
```

gattggccgt aacttttaga aaaaaatctt gttgatgata atgtgatttt ggggaggtca 2280
ttaattgctt tttctttttt aaatgtagac ttatataaat acctgtttgt atatagcttg 2340
agtaattgtg atatgattgt ataccactaa aatattgtta actattataa taaagtcaca 2400
gtaatggttt aaaaaagaaa aaagaaaaaa aaaaaaaaa 2439

<210> 2
<211> 529
<212> DNA
<213> Homo sapiens

<220>
<221>
<223> 1991226

<300>

<400> 2
ggacgctggc aaagcggggc agaccctgaa gactcactgc tcagcccagc gcccagatgt 60
ctgcaggtgg ctgagcccct tcatcctctc ctgctgcgtg tacttctgcc tctggattcc 120
cgaggaccag ctgtcctggt tcgctgccct ggtcaagtgc ctgcctgtcc tctgcctggc 180
tgggttcctg tgggtcatgt ccccaagcgg gggctacacc cagctcctcc agggagccct 240
tgtgtgctcg gctgtggggg acgcttgcct catctggccg gcagccttcg tccctggcat 300
ggccgccttt gccaccgccc acctcctcta cgtctgggcc ttcggcttct ctcccctgca 360
gcccggcctg ctgctgctca tcatcctggc ccctggcccc tactcagctt gtgctcagca 420
gctcgagccg gataggtctg cggtggagct atggctatct gatgcatctg tgcgggctgc 480
cacgtcgatg cgctgggcgg tgtctcacgt tctatggtgt gctggaact 529

<210> 3
<211> 2429
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 2404, 2413, 2423, 2425, 2426
<223> a or g or c or t, unknown, or other

<220>
<221>
<223> 253053

<300>

<400> 3
ttatgtgttt aacctataaa tattggggtc ttctgtctaa actggggtca ctgttgcatg 60
gaacattgtt cttaatagtt gagaattgct tttttgaaaa ttttcatgaa ggaattttgg 120
taatgacttt gcttgcaggt ttttttgggt gttttgagaa agtggcatgg aaacatgcag 180
tagttaatga gtttctcttg gtactgaaca ctattagaat atcattagtg atatttttc 240
tctttagagc atttttaatg caactagccc ctatatttta atgtaagagt tactctgcaa 300
tctaagcaaa gcacccaaca atggtaaatg ttttttaaaa atgcagaact aagatttttg 360
actctaaaga gagaaaatta caagggtgtt gccttatagc aaaccccttgg gacaatcctt 420
catgtgagca aagtgttgat cttaatattg gttgtctgtg gtgtgctttt ttgtactgta 480
aaaatatgtg gttcatgtct aactctgctg ttttattgtg gttgtggttc aagtttttaa 540
tgtttaaagt tgatgctgtt ttcagaagag ctttttacta atttatttgt cagtgttccc 600
tatttgttac ttaaccatga tcctccagat tttttggagt attcttttct aaccttaacc 660
ctgccaaacc ttgatccatt ttgacatttg ttatgcacta tttttatatc tctgtgagag 720
atttttccaa cagtcagcta ttttatggca cactttttt gactgatgac atctcctttg 780
ctatacctca atttttggaa tttagagaag aaatcagtag ttttgcaatg ttaattattt 840
agatattcaa tttcgcagat ttttaaactt tattttcata atttctgctt aatgtttaaa 900
attgaagagc cttttcatgt attaaataat gaacacaaat tatataatta aaataattgg 960
agatgttgaa aatcattttc ccttcttaaa cagaaataaa tatttggaat gaaggggaat 1020
gtactagaac acccttttg ccacgggtaa aaataacaga aatgtatggt ttgttttacc 1080
ttcatttctg tacaagtaaa gcttattagt ctaatgtttt gttcctttcc cacctcaccc 1140
ctacctcttt tgttttgttt tgtttttgcc ctttatgtac tacattctta ttttctaact 1200

tttaaacact gtattggagg ttttttttt taatttacag atcatattta ttttactatt 1260
tttgtagaaa attattaatt ttgattgtat ttttgtattt taaaagcttc ttcacttgtg 1320
ttccctaaat attcatattg ctgcccaaaa gtatgactgt ggaggaaaaa aaaatacttt 1380
aaaaatccac acttttttgtt aagaaggaaa catttagcat ttatatattt gtgtatggaa 1440
aacacttgat attttatccc tgttgcatct ggctgcacag agcctctcct caaagatgct 1500
acaaaacttg aatataacac attttggaag gctgactaac ctcgattctg tgttgtgatg 1560
tgcaatactg tttctaatgt ttgtataaaa aaaaacagtg taaacctttt taatgcaaat 1620
ttatttttt cattgcatat tttgcagatt ttatccacag tgtcattttt tactgtcaga 1680
aaagataccc cttttgtcat tgcaactatt ttttaaatcc agaaatcttt gtaactgatg 1740
taaatgattg tagttatttt gggatagtgt tttggctaac aaaagggaga gacttttttc 1800
atggcatatt tctatttttgt ttttttgggt tttattttat tttaatagtt agtaaaatac 1860
ttgggaataa tttttgcata ttcttgttca ttaatattat tttgtatttt tatgtggaaa 1920
tatataattt tatgacacta attgctaaag tttatttttat gttgaattat ttttggagct 1980
gaaatctttg taatattaaa gcaactagtt tctaattccc agtttctgta tagaatcgca 2040
caagtggttt atggagtgtt tggattgtaa ttataaatgg ttctttgata tgcaaattaa 2100
tattttcagt tgattttatt ttatattcct aatggggtgt taaagccgtt ttttattttt 2160
ttctaaataa aaagagaacc catgctttta tggacactag gtaaacacct tcagcttaaa 2220
tttttcgtta aatattttag tttattttat tgttatcttc caggtgtcta aatctccagt 2280
ctgtctgttg tactggtaat ttaactctgt aatggaatag tttgctgcca actatttata 2340
ttaagtaatt tttaaatatt tgtaatattg ttgactgact aataaactat taagttattg 2400
gaangaaaaa canaaaaaaa aanannaac 2429

<210> 4
<211> 944
<212> DNA
<213> Homo sapiens

<220>
<221>
<223> 2009569

<300>

<400> 4
gcttgaaggc gggaaaatac ggaatggagc gacagggagt ggacgtgccc catgtgaaat 60
gcaaagacca ggaaccgcag cccttggggg agagcaagga gcatccgcgg tgggaagaga 120
actgcgagga ggaagctggt ggagggccag ctagtgccag ttgccagctg acggtcctgg 180
aagggaagtc gggactctac ttctcctctc tggactcaag cattgacatc ctgcagaaga 240
gagcccagga gctgatcgaa aacatcaaca agagccggca aaaggaccat gcactcatga 300
ccaacttcag gaacagcctg aagaccaagg tttcggatct gacagagaaa ttagaggaga 360
ggatctatca gatttataat gaccacaaca agatcatcca ggaaaagctc caagagttca 420
cccagaaaat ggcaaagatc agccatttgg agacagagct caaacaagtc tgccacagcg 480
tggagactgt gtacaaagac ctgtgtctcc agcctgagca gagcctaaga ctcagatggg 540
ggccagacca ctctagggga aagtccccac cacgtcccgg caactcacag cccccagacg 600
tgttcgtttc ttctgtggct gaaactactt ctcaggccac tgcttcagaa gtacagacca 660
acagagatgg tgaatgctga cagctgccgg gagactcacg ccttagtgac agtctccagg 720
agaagactgt gaggccacca tttgggccac actgagaaat tgtttttcat ggttctataa 780
tgcatcttgg cagaaaaaaa caaaaaccca aagctccttg tgctgaactc ccaaaatgta 840
gcaagtccag ccctctccat aggcccaggc ttcgtgctcc ccacccttgg caagttctcc 900
ccacccccag ccccacagtt tattaaatgt ttgattttca aaaa 944

<210> 5
<211> 404
<212> DNA
<213> Homo sapiens

<220>
<221>
<223> 1642580

<300>

<400> 5
ctcaagaccc agcagtggga cagccagaca gacggcacga tggcactgag ctcccagatc 60

tgggccgctt gcctcctgct cctcctcctc ctcgccagcc tgaccagtgg ctctgttttc 120
ccacaacaga cgggacaact tgcagagctg caaccccagg acagagctgg agccagggcc 180
agctggatgc ccatgttcca gaggcgaagg aggcgagaca cccacttccc catctgcatt 240
ttctgctgcg gctgctgtca tcgatcaaag tgtgggatgt gctgcaagac gtagaaccta 300
cctgccctgc ccccgtcccc tcccttcctt atttattcct gctgccccag aacataggtc 360
ttggaataaa atggctggtt cttttgtttt ccaaaaaaaa aaaa 404

<210> 6
<211> 587
<212> PRT
<213> Homo sapiens

<220>
<221>
<223> 1851405

<300>

<400> 6
Met Ser Asn Val Met Gly Asn Thr Thr Trp Thr Thr Ser Gly Leu
1 5 10 15
Lys Ser Gln Gly Arg Leu Ser Val Gly Ser Asn Arg Asp Arg Glu
20 25 30
Ile Ser Met Ser Val Gly Leu Gly Arg Ser Gln Leu Asp Ser Lys
35 40 45
Gly Gly Val Val Gly Gly Thr Ile Asp Val Asn Ala Leu Glu Met
50 55 60
Val Ala His Ile Ser Glu His Pro Asn Gln Gln Pro Ser His Lys
65 70 75
Ile Gln Ile Thr Met Gly Ser Thr Glu Ala Arg Val Asp Tyr Met
80 85 90
Gly Ser Ser Ile Leu Met Gly Ile Phe Ser Asn Ala Asp Leu Lys
95 100 105
Leu Gln Asp Glu Trp Lys Val Asn Leu Tyr Asn Thr Leu Asp Ser
110 115 120
Ser Ile Thr Asp Lys Ser Glu Ile Phe Val His Gly Asp Leu Lys
125 130 135
Trp Asp Ile Phe Gln Val Met Ile Ser Arg Ser Thr Thr Pro Asp
140 145 150
Leu Ile Lys Ile Gly Met Lys Leu Gln Glu Phe Phe Thr Gln Gln
155 160 165
Phe Asp Thr Ser Lys Arg Ala Leu Ser Thr Trp Gly Pro Val Pro
170 175 180
Tyr Leu Pro Pro Lys Thr Met Thr Ser Asn Leu Glu Lys Ser Ser
185 190 195
Gln Glu Gln Leu Leu Asp Ala Ala His His Arg His Trp Pro Gly
200 205 210
Val Leu Lys Val Val Ser Gly Cys His Ile Ser Leu Phe Gln Ile
215 220 225
Pro Leu Pro Glu Asp Gly Met Gln Phe Gly Gly Ser Met Ser Leu
230 235 240
His Gly Asn His Met Thr Leu Ala Cys Phe His Gly Pro Asn Phe
245 250 255
Arg Ser Lys Ser Trp Ala Leu Phe His Leu Glu Glu Pro Asn Ile
260 265 270
Ala Phe Trp Thr Glu Ala Gln Lys Ile Trp Glu Asp Gly Ser Ser
275 280 285
Asp His Ser Thr Tyr Ile Val Gln Thr Leu Asp Phe His Leu Gly
290 295 300
His Asn Thr Met Val Thr Lys Pro Cys Gly Ala Leu Glu Ser Pro
305 310 315
Met Ala Thr Ile Thr Lys Ile Thr Arg Arg Arg His Glu Asn Pro
320 325 330
Pro His Gly Val Ala Ser Val Lys Glu Trp Phe Asn Tyr Val Thr
335 340 345

Ala Thr Arg Asn Glu Glu Leu Asn Leu Leu Arg Asn Val Asp Ala
350 355 360
Asn Asn Thr Glu Asn Ser Thr Thr Val Lys Asn Ser Ser Leu Leu
365 370 375
Ser Gly Phe Arg Gly Gly Ser Ser Tyr Asn His Glu Thr Glu Thr
380 385 390
Ile Phe Ala Leu Pro Arg Met Gln Leu Asp Phe Lys Ser Ile His
395 400 405
Val Gln Glu Pro Gln Glu Pro Ser Leu Gln Asp Ala Ser Leu Lys
410 415 420
Pro Lys Val Glu Cys Ser Val Val Thr Glu Phe Thr Asp His Ile
425 430 435
Cys Val Thr Met Asp Ala Glu Leu Ile Met Phe Leu His Asp Leu
440 445 450
Val Ser Ala Tyr Leu Lys Glu Lys Glu Lys Ala Ile Phe Pro Pro
455 460 465
Arg Ile Leu Ser Thr Arg Pro Gly Gln Lys Ser Pro Ile Ile Ile
470 475 480
His Asp Asp Asn Ser Ser Asp Lys Asp Arg Glu Asp Ser Ile Thr
485 490 495
Tyr Thr Thr Val Asp Trp Arg Asp Phe Met Cys Asn Thr Trp His
500 505 510
Leu Glu Pro Thr Leu Arg Leu Ile Ser Trp Thr Gly Arg Lys Ile
515 520 525
Asp Pro Val Gly Val Asp Tyr Ile Leu Gln Lys Leu Gly Phe His
530 535 540
His Ala Arg Thr Thr Ile Pro Lys Trp Leu Gln Arg Gly Val Met
545 550 555
Asp Pro Leu Asp Lys Val Leu Ser Val Leu Ile Lys Lys Leu Gly
560 565 570
Thr Ala Leu Gln Asp Glu Lys Glu Lys Lys Gly Lys Asp Lys Glu
575 580 585
Glu His

<210> 7
<211> 218
<212> PRT
<213> Homo sapiens

<220>
<221>
<223> 2009569

<300>

<400> 7
Met Glu Arg Gln Gly Val Asp Val Pro His Val Lys Cys Lys Asp
1 5 10 15
Gln Glu Pro Gln Pro Leu Gly Glu Ser Lys Glu His Pro Arg Trp
20 25 30
Glu Glu Asn Cys Glu Glu Glu Ala Gly Gly Gly Pro Ala Ser Ala
35 40 45
Ser Cys Gln Leu Thr Val Leu Glu Gly Lys Ser Gly Leu Tyr Phe
50 55 60
Ser Ser Leu Asp Ser Ser Ile Asp Ile Leu Gln Lys Arg Ala Gln
65 70 75
Glu Leu Ile Glu Asn Ile Asn Lys Ser Arg Gln Lys Asp His Ala
80 85 90
Leu Met Thr Asn Phe Arg Asn Ser Leu Lys Thr Lys Val Ser Asp
95 100 105
Leu Thr Glu Lys Leu Glu Glu Arg Ile Tyr Gln Ile Tyr Asn Asp
110 115 120
His Asn Lys Ile Ile Gln Glu Lys Leu Gln Glu Phe Thr Gln Lys
125 130 135
Met Ala Lys Ile Ser His Leu Glu Thr Glu Leu Lys Gln Val Cys

```
                  140                 145                 150
His Ser Val Glu Thr Val Tyr Lys Asp Leu Cys Leu Gln Pro Glu
                  155                 160                 165
Gln Ser Leu Arg Leu Arg Trp Gly Pro Asp His Ser Arg Gly Lys
                  170                 175                 180
Ser Pro Pro Arg Pro Gly Asn Ser Gln Pro Pro Asp Val Phe Val
                  185                 190                 195
Ser Ser Val Ala Glu Thr Thr Ser Gln Ala Thr Ala Ser Glu Val
                  200                 205                 210
Gln Thr Asn Arg Asp Gly Glu Cys
                  215
```

<210> 8
<211> 84
<212> PRT
<213> Homo sapiens

<220>
<221>
<223> 1642580

<300>

<400> 8

```
Met Ala Leu Ser Ser Gln Ile Trp Ala Ala Cys Leu Leu Leu Leu
  1               5                  10                  15
Leu Leu Leu Ala Ser Leu Thr Ser Gly Ser Val Phe Pro Gln Gln
                  20                  25                  30
Thr Gly Gln Leu Ala Glu Leu Gln Pro Gln Asp Arg Ala Gly Ala
                  35                  40                  45
Arg Ala Ser Trp Met Pro Met Phe Gln Arg Arg Arg Arg Arg Asp
                  50                  55                  60
Thr His Phe Pro Ile Cys Ile Phe Cys Cys Gly Cys Cys His Arg
                  65                  70                  75
Ser Lys Cys Gly Met Cys Cys Lys Thr
                  80
```

<210> 9
<211> 644
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 44
<223> a or g or c or t, unknown, or other

<220>
<221>
<223> 700062690

<300>

<400> 9

```
ttgatctagt gtcagcatat ctgaaagaaa aggaaaaggc catntttcca cctcggattt  60
tatccactcg accaggacaa aagtgtccag ttattataca ggatgacaat tcctctgaca 120
gagacagaga agacggcatc acttacacca ctgtggactg gcgagacttt atgtgtaaca 180
catggcatct ggagcctact cttagattaa tttcttggac tggaagaaag attgatccag 240
taggtgttga ttacattctt caaaaattgg gctttcacca tgccaggact actattccca 300
aatggcttca gagaggagtg atggacccac tggataaggt gctatccgtt ctcatcaaaa 360
aactcgggac tgccctgcag gatgagaagg agaagaaagg aaaagacaaa gaagaacact 420
agaatagtaa ctccatctgt gaagtacctg actctgtccg ttaagttctg gcactgcaat 480
tttttcagca taactttaaa tggtacttca aaataattct aattttgtgg ccattattga 540
aacagtttgt aagttaatct gtttattcta gctctaccat tctgcataca gtaaagtatt 600
```

gcatgaaatg tacattttgt agaaaaacta aattaaaaga tata 644

<210> 10
<211> 449
<212> DNA
<213> Rattus norvegicus

<220>
<221>
<223> 700140450

<300>

<400> 10
gttcctaaag agaagtcact caacatggat gctatgaaaa gagggactgt ccctggagat 60
actcttttca gaccaatatc cacaggtccg caggtggctg gccccttca tcgttgcctg 120
ctccctctac ttcctcctct ggattcctga ggaccagcca tcttgggtca gtgccctggt 180
caagtgccag cccattctct gcctggtttt gttcctgtgg gctgtggctc ctggtgggag 240
ctacacctgg ctcctgcagg gagctcttac atgttctgct gttggagatg cctgcctcat 300
ctggcctgaa gctttctttt atggcatggc agtcttctct gttgcccact tactctacct 360
tctgacatac tacagcttcc tcctgctac 449

<210> 11
<211> 259
<212> DNA
<213> Rattus norvegicus

<220>
<221>
<223> 700367683

<300>

<400> 11
gccgtttttt attttttct aaataaaaag agaacccatg cttttatgga cactaggtaa 60
aacgccttca gcttaaaatt tttgttaaat aatttagttt attttattgt tatcttccag 120
gtgtctaaat ctccagtctg tctgttgtac tggtaattta actctgtaat ggaatagttt 180
gctgccaact atttatatta agtcattttt aaatatttgt aatattgttg actgactaat 240
aaactattaa gttattggc 259

<210> 12
<211> 261
<212> DNA
<213> Rattus norvegicus

<220>
<221>
<223> 701316810

<300>

<400> 12
gtctcagatc tgacagaaaa gttggaggaa aggatgtacc aggtatacag ccaccacagc 60
aaaatcattc aggaaagact acaagaattt acccagaaga tggcaaagat cagtcatctg 120
gaaatggagc tcaaacaagt ttgccaaact gtggaaacta tgtacaagga cctatgtgtc 180
cagtctgagg taagaatttg ggggtgtatc ctgtgagaaa ggagaccgga aagtaaaaca 240
tgaatgggaa ctttgacttc t 261

<210> 13
<211> 406
<212> DNA

<213> Rattus norvegicus

<220>
<221>
<223> 700139271

<300>

<400> 13
tcaagtcctt ggactacact gcaggacaga aggcaagatg gcactaagca ctcggatcca 60
ggctgcctgt ctcctgcttc tcctcctggc cagcctgagc agcggtgcct atctccggca 120
acagacgaga cagactacgg ctctgcagcc ttggcatggg gcagaaagca agactgatga 180
cagtgcgctg ctgatgctga agcgaaggaa gcgagacacc aacttcccca tatgcctctt 240
ctgctgtaaa tgctgtaaga attcctcctg tggtctctgt tgcataacat agagagccaa 300
gagccttgtc ctgacctctc aacacactgc ctcccctccg ccccattatt tattcctgtc 360
ctaccccagc aatgaccttg aaataaaaat gatgatttta ttttca 406

<210> 14
<211> 84
<212> PRT
<213> Rattus norvegicus

<220>
<221>
<223> 700139271

<300>

<400> 14
Met Ala Leu Ser Thr Arg Ile Gln Ala Ala Cys Leu Leu Leu Leu
1 5 10 15
Leu Leu Ala Ser Leu Ser Ser Gly Ala Tyr Leu Arg Gln Gln Thr
20 25 30
Arg Gln Thr Thr Ala Leu Gln Pro Trp His Gly Ala Glu Ser Lys
35 40 45
Thr Asp Asp Ser Ala Leu Leu Met Leu Lys Arg Arg Lys Arg Asp
50 55 60
Thr Asn Phe Pro Ile Cys Leu Phe Cys Cys Lys Cys Cys Lys Asn
65 70 75
Ser Ser Cys Gly Leu Cys Cys Ile Thr
80